Combined Channel Coding and Space-Block Coding In a Multi-Antenna Arrangement

Reference to Related Applications

5   This application is related to:

- U.S. Application 09/074,224, filed May 7, 1998, titled "Transmitter Diversity Technique for Wireless Communications," which is based U.S. Provisional Applications No. 60/059,016 filed Sep. 16, 1997; 60/059,219 filed Sep. 18, 1997; and 60/063,780 filed Oct. 31, 1997;

10 - U.S. Application 09/149, 163 filed Sep. 4, 1999, titled "Combined Interference Cancellation and Maximum Likelihood Decoding of Block Space-Time Coded Modulation," which is based on U.S. Provisional Application Nl. 60/061,145, filed October 6, 1997; and   09/114,838

- U.S. Application 09/___, filed July 10, 1998, titled "Combined Array Processing and Space-Time Coding," which is based on U.S. Provisional Application No. 60/052,689, filed July 17, 1997.

This application claims the benefit of U.S. Provisional Application No. 60/099,212, filed September 4, 1998.

Background of the Invention

20   This invention relates to wireless communication and, more particularly, to techniques for effective wireless communication in the presence of fading, co-channel interference, and other degradations.

Physical limitations of wireless channels present a fundamental technical challenge to reliable communication. Bandwidth limitations, propagation loss, time variance, noise, 25 interference, and multipath fading make the wireless channel a narrow "pipe" that does not easily accommodate the flow of data. Further challenges come from power limitation, size, and speed of devices that are employed within portable wireless apparatus.

Employing multiple transmit antennas at both the base and the remote stations increases the capacity of wireless channels, and information theory provides a measure of 30 this increase. The standard approach to exploit this capacity is linear processing at the receiver, as described, for example, by J. Winters, J. Salz and R. D. Gitlin, in "The impact

of antenna diversity an the capacity of wireless communication systems" *IEEE Trans. Communications*, Vol. 42. No. 2/3/4, pp. 1740-1751, Feb/March/April 1994. Transmit diversity has been explored by Wittneben in "Base station modulation diversity for digital SIMULCAST," *Proc. IEEE'VTC*, pp. 505-511, May 1993, and by Seshadri and Winters in

5   "Two signaling schemes for improving the error performance of frequency-division-duplex (FDD) transmission systems using transmitter antenna diversity," *International Journal of Wireless Information Networks*, Vol. 1, No. 1, 1994. The Wittneben and Seshadri et al papers approach transmit diversity from a signal processing point of view.

Space-time codes combine signal processing at the receiver with coding techniques

10   appropriate to multiple transmit antennas. See, for example, V. Tarokh, N. Seshadri, and A. R. Calderbank in "Space-Time Codes For High Data Rate Wireless Communication: Performance Analysis and Code Construction," *IEEE Trans. Info. Theory*, Vol. 44, No. *2*, pp. 744-765, March 1998. The space-time approach provides significant gain over the aforementioned prior art. Specific space-time codes designed for 2-4 transmitting antennas

15   perform well in slowly varying fading environments (such as indoor transmission) and come with in 2-3 dB of the theoretical outage capacity. Outage capacity is described, for example, by J. Foschini, Jr. and M. J. Gans, "On limits of wireless communication in a fading environment, when using multiple antennas," *Wireless Personal Communication*, Vol. 6, No. 3, pp. 311-335, March 1998. The bandwidth efficiency of the codes described

20   in the Tarokh et al paper is about 3-4 times that of the current systems. The most important contributor to the improved performance is diversity, which may be thought of as providing a number of replicas of the transmitted signal to the receiver, with some replicas being less attenuated by fading. The space-time codes presented in the Tarokh et al paper provide an optimized trade-off between constellation size, data rate, diversity gain,

25   and trellis complexity

When the number of transmit-antennas is fixed, decoding complexity (measured, for example, by the number of trellis states in the decoder) increases exponentially with transmission rate. This can be refined to some extent by designing space-time codes with a multi-level structure and adopting multi-stage decoding, as described by Tarokh et al. For

30   a moderate number of transmit antennas (3-6), this method provides higher data rates while reducing the decoding complexity. However, there is a penalty to be paid for simplified

decoding. Multi-stage decoding is sub-optimal, due in part to magnification in the error coefficient, and this performance penalty means that alternative solutions are needed for achieving very high data rates.

In order to achieve very high data rates on narrowband wireless channels, many
5 antennas are needed at both the transmitter and the receiver. Consider a wireless communication system employing n transmit and m receive antennas, where the subchannel between each transmit and receive antenna is quasi-static Rayleigh, flat, and independent of others. If n is fixed, then the capacity increases logarithmically with m. On the other hand, if m is fixed then intuition suggests that there must come a point where
10 adding more transmit antennas will not make much difference. Indeed, this can be seen in the mathematics of outage capacity, shown by Foschini and Gans in the aforementioned paper. Thus, it turns out that in the presence of one receive antenna little can be gained in terms of outage capacity by using more than 4 transmit antennas. A similar argument shows that if there are two receive antennas, using 6 transmitting antennas provides almost
15 all of the capacity increase that can be obtained.

If n increases and $m \geq n$, then information theory shows that the capacity of the system increases at least linearly as a function of n. Thus, it makes sense to increase the number of antennas at both receiver and transmitter in order to obtain higher capacities. Employing many antennas at both transmitter and receiver creates a multiple-input
20 multiple-output system where the number of degrees of freedom is given by the product of the number of transmit and receive antennas.

Foschini considered such a system in "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal,* Vol. 1, No. 2, Autumn 1996. He proposed a multi-layered structure
25 that, in principle, can achieve a tight lower bound on the capacity. If n transmit and n receive antennas are used, then at the receiver the transmitted signal from transmit antenna 1 is treated as the desired signal, while the signals transmitted from other transmit antennas are treated as interference. Linear processing is then used to suppress the interfering signals using n receive-antennas, providing a diversity gain of one. Once the signal
30 transmitted from antenna 1 is correctly detected, the signal transmitted from antenna 2 is treated as the desired signal while the signals transmitted from transmit antennas 3, 4, ..., n

are treated as interference. The contribution of the signal transmitted from antenna 1, now that it has been detected, is subtracted from the signal received at receiver antennas 1 through n. Thereafter, detection of the signal transmitted by antenna 2 proceeds with linear processing that is applied to suppress the interfering signals from antennas 3 through n.

5   This provides a diversity gain of two. This process is repeated until all the transmitted signals are detected. Clearly, the worst case diversity in this architecture is one. For such a system, long frames of data combined with powerful coding techniques are needed to achieve the lower bound on outage capacity.

In the aforementioned 09/.... application, an arrangement is disclosed that provides

10   enhanced performance which is realized by employing a perspective that combines array signal processing with channel coding. Specifically, antennas at the transmitter are partitioned into small groups, and individual space-time codes are used to transmit information from each group of antennas. At the receiver, an individual space-time code is decoded by a linear array processing technique that suppresses signals transmitted by other

15   groups of antennas by treating them as interference. The contribution of the decoded signal to other received signal is then subtracted from those received signals. What results is a simple receiver structure that provides diversity and coding gain over uncoded systems with a given diversity gain. This combination of array processing at the receiver and coding techniques for multiple transmit antennas provides reliable and very high data rate

20   communication over wireless channels. One advantage of the group interference suppression method over the Foschini architecture is that the number of receive antennas can be less than the number of transmit antennas.

In the aforementioned 09/149,163 application, an arrangement is disclosed wherein K synchronized terminal units transmit on N antennas to a base station having $M \geq K$

25   antennas. An improvement is achieved by employing both interference cancellation (IC) and maximum likelihood (ML) decoding. More specifically, space-time block coding is employed in transmitters that employ N transmit antennas each, and the signals are received in a receiver that employs M receiving antennas. By exploiting the structure of the space-time block code, K-1 interfering transmitting units are cancelled at the receiver,

30   regardless of the number of transmitting antennas, N, when decoding the signals transmitted by a given mobile unit. Also disclosed is an arrangement where signals of a

first terminal unit are decoded first, and the resulting decoded signals are employed to cancel their contribution to the signals received at the base station antennas while decoding the signals of the remaining K-1 terminal units. The process is repeated among the remaining K-1 terminal units.

5

Summary of the Invention

Enhanced performance is achieved by combining channel coding with the space-time coding principles disclosed in the '163 applications. More specifically, with K synchronized terminal units transmitting on N antennas to a base station having $M \geq K$ receive antennas, increased system capacity and improved performance are attained by using a concatenated coding scheme where the inner code is a space-time block code and the outer code is a conventional channel error correcting code. That is, information symbols are first encoded using a conventional channel code. The channel code is then encoded signal is then encoded using a space-time block code, and transmitted over N antennas. At the receiver, the inner space-time block code is used to suppress interference from the other co-channel terminals and soft decisions are made about the transmitted symbols. The channel decoding that follows makes the hard decisions about the transmitted symbols.

Increased data rate is achieved by, effectively, splitting the incoming data rate into multiple channels, and each channel is transmitted over its own terminal. Viewed another way, information symbols from a transmitting terminal is split into L parallel streams. Stream l is encoded using a channel code with rate R_l and then coded with a space-time block encoder with N transmitting antennas. Advantageously, the coding rates are chosen such that $R_1 > R_2 > \cdots > R_L$.

Brief Description of the Drawing

FIG. 1 depicts an arrangement that, illustratively, includes a base station 30 with four antennas, a terminal units 20 with two antennas, and a terminal unit 20 with two antennas; and

FIG. 2 illustrates a terminal unit that splits the input signal into two streams, and each stream is transmitted over a separate two-antenna arrangement.

Detailed Description

FIG. 1 illustrates a unit 10 that employs a space-time block coding unit 13 that is followed by a conventional constellation mapper and pulse shaping circuit 16. The output of circuit 16 is fed to two transmitting antennas 11 and 12. The input symbols to the space-time block encoder are divided into groups of two symbols each, and at a given symbol period, the two symbols in each group $\{c_1, c_2\}$ are transmitted simultaneously from the two antennas. The signal transmitted from antenna 11 is c_1 and the signal transmitted from antenna 12 is c_2. In the next symbol period, the signal $-c_2{}^*$ is transmitted from antenna 11 and the signal $c_1{}^*$ is transmitted from antenna 12.

In receiver 20, signals are received by antennas 21 and 22 and are applied to detector 25. Channel estimators 23 and 24 operate on the incoming signal of antennas 21 and 24, respectively, in a conventional manner to develop estimates of the channel parameters. Those estimates are applied to detector 25. In the mathematical development of the algorithms disclosed herein, it is assumed that the channel from each of the two transmit antennas remains fixed over a two consecutive symbol periods. That is,

$$h_i(nT) = h_i((n+1)T), \quad i = 1,2. \tag{1}$$

To ascertain the channel characteristics, the transmitter carries out a calibration session, during which pilot signals or tones are transmitted. It is the signals that are received during the calibration session that are utilized by the channel estimator circuits 23 and 24, which are well known circuits.

Maximum Likelihood Detection

The received signals at antenna 21 can be expressed as

$$r_1 = h_1 c_1 + h_2 c_2 + \eta_1 \tag{2}$$

$$r_2 = -h_1 c_2^* + h_2 c_1^* + \eta_2, \tag{3}$$

where r_1 and r_2 are the received signals over two consecutive symbol periods, h_1 denotes the fading channel between transmit antenna 11 and receive antenna 21, h_2 denotes channel between transmit antenna 12 and receive antenna 21, and η_1 and η_2 are noise terms, which are assumed to be complex Gaussian random variables with zero mean and

power spectral density $N_0/2$ per dimension. Defining the vectors $\mathbf{r} = [r_1 r_2 {}^*]^T$, $\mathbf{c} = [c_1 c_2]^T$,

and $\eta = [\eta_1 \eta_2 {}^*]^T$, equations (2) and (3) can be rewritten in a matrix form as

$$\mathbf{r} = \mathbf{H} \cdot \mathbf{c} + \eta, \qquad (4)$$

where the channel matrix $\mathbf{H}$ is defined as

$$\mathbf{H} = \begin{bmatrix} h_1 & h_2 \\ h_2{}^* & -h_1{}^* \end{bmatrix}. \qquad (5)$$

The vector η is a complex Gaussian random vector with zero mean and covariance $N_0 \cdot \mathbf{I}$.

Defining C as the set of all possible symbol pairs $\mathbf{c} = \{c_1, c_2\}$, and assuming that all

symbol pairs are equi-probable, it can be easily shown that the optimum maximum

likelihood (ML) decoder selects from C the symbol pair $\hat{\mathbf{c}}$ that minimizes the expression

$\|\mathbf{r} - \mathbf{H} \cdot \hat{\mathbf{c}}\|^2$. This can be written as

$$\hat{\mathbf{c}} = \arg \min_{\hat{\mathbf{c}} \in C} \|\mathbf{r} - \mathbf{H} \cdot \hat{\mathbf{c}}\|^2. \qquad (6)$$

It was shown by S. Alamouti in "Space Block Coding: A simple Transmitter

Diversity Scheme for wireless Communications," submitted to IEEE JSAC, September

1997, that the diversity order of the above space-time block code is equivalent to that of a

two branch maximal ratio receive combining (MRRC). Alamouti also showed that,

because of the orthogonality of the matrix $\mathbf{H}$, this decoding rule decomposed into two

separate decoding rules for c_1 and c_2. The *uncertainty*, Δ_c, of the decoded symbols $\hat{\mathbf{c}}$ is

defined as

$$\Delta_c = \|\mathbf{r} - \mathbf{H} \cdot \hat{\mathbf{c}}\|^2. \qquad (7)$$

The maximum likelihood (ML) rule of equation (6) can be simplified by realizing

that the channel matrix $\mathbf{H}$ is orthogonal; that is, that $\mathbf{H}^*\mathbf{H} = (|h_1|^2 + |h_2|^2)\mathbf{I}$. This yields a

modified received vector,

$$\tilde{\mathbf{r}} = \mathbf{H}^*\mathbf{r} = (|h_1|^2 + |h_2|^2) \cdot \mathbf{c} + \tilde{\eta}, \qquad (8)$$

where $\tilde{\eta} = \mathbf{H}^*\eta$. This leads to

$$\hat{\mathbf{c}} = \arg \min_{\hat{\mathbf{c}} \in C} \|\tilde{\mathbf{r}} - (|h_1|^2 + |h_2|^2) \cdot \hat{\mathbf{c}}\|^2. \qquad (9)$$

Thus it follows that by using a simple linear combining, the decoding rule of equation (9)

7

reduces to two separate, and much simpler, decoding rules for c_1 and c_2. When using a signaling constellation with 2b constellation points, this reduces the number of decoding metrics that have to be computer for ML decoding from 2^{2b} to 2×2^b.

When receiver 20 uses M receiver antennas, the received vector at antenna m is

$$\mathbf{r}_m = \mathbf{H}_m \cdot \mathbf{c} + \eta_m, \tag{10}$$

where the channel matrix $\mathbf{H}_m$ is defined as

$$\mathbf{H}_m = \begin{bmatrix} h_{1m} & h_{2m} \\ h_{2m}^* & -h_{1m}^* \end{bmatrix}. \tag{11}$$

In this case, the optimum ML decoding rule is

$$\hat{\mathbf{c}} = \arg \min_{\hat{\mathbf{c}} \in C} \sum_{m=1}^{M} \left\| \mathbf{r}_m - \mathbf{H}_m \cdot \hat{\mathbf{c}} \right\|^2, \tag{12}$$

and the corresponding *uncertainty*, Δ_c, of the decoded symbols $\hat{\mathbf{c}}$ is defined by

$$\Delta_c = \sum_{m=1}^{M} \left\| \mathbf{r}_m - \mathbf{H}_m \cdot \hat{\mathbf{c}} \right\|^2. \tag{13}$$

As before, in the case of M receive antennas, the decoding rule can be simplified by premultiplying the received signal by $\mathbf{H}_m^*$.

As indicated above, FIG. 1 shows two terminal units 10 and 30, and the issue that needs to be addressed is the detection performance at the base station receiver when the two terminal units synchronously transmit *over the same time and frequency channel.*

In the notation below, g_{11} denotes the fading channel between transmit antenna 31 and receive antenna 21, g_{12} denotes the channel between antenna 31 and antenna 22, g_{21} denotes the channel between antenna 32 and antenna 21, and g_{22} denotes the channel between antenna 32 and antenna 22. Also, $\{c_1, c_2\}$ and $\{s_1, s_2\}$ denote the two symbols transmitted from terminal units 10 and 30, respectively.

At receiver 20, the received signals over two consecutive symbol periods at receive antenna 21, r_{11} and r_{12}, are

$$r_{11} = h_{11}c_1 + h_{21}c_2 + g_{11}s_1 + g_{21}s_2 + \eta_{11} \tag{14}$$

$$r_{12} = -h_{11}c_2^* + h_{21}c_1^* - g_{11}s_2^* + g_{21}s_1^* + \eta_{12}. \tag{15}$$

Defining $\mathbf{r}_1 = [r_{11} r_{12}^*]^T$, $\mathbf{c} = [c_1 \ c_2]^T$, $\mathbf{s} = [s_1 \ s_2]^T$, and $\mathbf{n}_1 = [\eta_{11} \ \eta_{12}^*]^T$ equations (14) and

(15) can be rewritten in matrix form as

$$\mathbf{r}_1 = \mathbf{H}_1 \cdot \mathbf{c} + \mathbf{G}_1 \cdot \mathbf{s} + \mathbf{n}_1, \tag{16}$$

where the channel matrices $\mathbf{H}_1$ and $\mathbf{G}_1$ between the transmitter units 10 and 30 and receive antenna 21 are given by

$$\mathbf{H}_1 = \begin{bmatrix} h_{11} & h_{21} \\ h_{21}^* & -h_{11}^* \end{bmatrix}, \text{ and } \mathbf{G}_1 = \begin{bmatrix} g_{11} & g_{21} \\ g_{21}^* & -g_{11}^* \end{bmatrix}. \tag{15}$$

The vector $\mathbf{n}_1 = [\eta_{11}\eta_{12}*]^T$ is a complex Gaussian random vector with zero mean and covariance $N_0 \cdot \mathbf{I}$. Similarly, the received signals over two consecutive symbol periods at receive antenna 22, r_{21} and r_{22}, are

$$r_{21} = h_{12}c_1 + h_{22}c_2 + g_{12}s_1 + g_{22}s_2 + \eta_{21} \tag{16}$$

$$r_{22} = -h_{12}c_2^* + h_{22}c_1^* - g_{12}s_2^* + g_{22}s_1^* + \eta_{22}. \tag{17}$$

In a similar fashion, defining $\mathbf{r}_2 = [r_{21}r_{22}*]^T$ and $\mathbf{n}_2 = [\eta_{21}\eta_{22}*]^T$ equations (16) and (17) can be rewritten as

$$\mathbf{r}_2 = \mathbf{H}_2 \cdot \mathbf{c} + \mathbf{G}_2 \cdot \mathbf{s} + \mathbf{n}_2, \tag{18}$$

where the channel matrices $\mathbf{H}_2$ and are given by

$$\mathbf{H}_2 = \begin{bmatrix} h_{12} & h_{22} \\ h_{22}^* & -h_{12}^* \end{bmatrix}, \text{ and } \mathbf{G}_2 = \begin{bmatrix} g_{12} & g_{22} \\ g_{22}^* & -g_{12}^* \end{bmatrix}. \tag{19}$$

Equations (14) and (18) can be combined to yield the matrix form

$$\mathbf{r} = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{H}_1 & \mathbf{G}_1 \\ \mathbf{H}_2 & \mathbf{G}_2 \end{bmatrix} \begin{bmatrix} \mathbf{c} \\ \mathbf{s} \end{bmatrix} + \begin{bmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \end{bmatrix}. \tag{20}$$

Minimum Mean-Squared Error Interference Cancellation (MMSE IC)

When seeking to detect and decode signals $\{c_1, c_2\}$ by minimizing a mean-squared error criterion, the goal is find a linear combination of the received signals such that the mean-squared error in detecting the signals $\{c_1, c_2\}$ is minimized. In general terms, this can be expressed by an error cost function that is to be minimized, such as the function

$$J(\alpha, \beta) = \left\| \sum_{i=1}^{4} \alpha_i^* r_i - (\beta_i^* c_1 + \beta_i^* c_2) \right\|^2 = \left\| \alpha^* \cdot \mathbf{r} - \beta^* \cdot \mathbf{c} \right\|^2, \tag{21}$$

where $\mathbf{r} = \begin{bmatrix} r_1 & r_2 & r_3 & r_4 \end{bmatrix}^T = \begin{bmatrix} r_{11} & r_{12} & r_{21} & r_{22} \end{bmatrix}^T.$

One may note that a minimum is certainly reached when both α and β are equal to 0, but that, of course, is not desired. Therefore, either β_1 or β_2 is set to 1.

~~When β_1 is set to 1, we get the following minimization criterion from equation (40)~~

$$J_1(\alpha_1, \beta_1) = \left\| \sum_{i=1}^{5} \alpha_{1i}^* r_{1i} - c_1 \right\|^2 = \left\| \widetilde{\alpha}_1^* \widetilde{r}_1 - c_1 \right\|^2 = J_1(\widetilde{\alpha}_1), \qquad (22)$$

where $\widetilde{\alpha}_1 = [\alpha_{11}, \alpha_{12}, \alpha_{13}, \alpha_{14}, -\beta_2] = [\alpha_1 \quad -\beta_2]$ and $\widetilde{r}_1 = [r^T \quad c_2]^T$. From this it can be seen that

$$\widetilde{r}_1 = \begin{bmatrix} \mathbf{H} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} \widetilde{c} \\ c_2 \end{bmatrix} + \begin{bmatrix} \mathbf{n} \\ 0 \end{bmatrix} = \mathbf{R} \cdot \mathbf{d}_1 + \widetilde{\eta}, \qquad (23)$$

where $\mathbf{0} = [0 \quad 0 \quad 0 \quad 0]^T$.

What is needed is to select $\widetilde{\alpha}_1$ so that the expected value of the expression in equation (22) is minimized. That is, select $\widetilde{\alpha}_1$ to minimize

$$\begin{aligned} E\{J_1(\widetilde{\alpha}_1)\} = E\{J_1(\widetilde{\alpha}_1)\} &= E\{(\widetilde{\alpha}_1^* \widetilde{r}_1 - c_1)(\widetilde{\alpha}_1^* \widetilde{r}_1 - c_1)^*\} \\ &= \widetilde{\alpha}_1^* E\{\widetilde{r}_1 \widetilde{r}_1^*\} \widetilde{\alpha}_1 + E\{c_1 c_1^*\} - E\{\widetilde{r}_1 c_1^*\} \widetilde{\alpha}_1^* - E\{c_1 \widetilde{r}_1^*\} \widetilde{\alpha}_1 \end{aligned} \qquad (24)$$

Taking the partial derivative with respect to $\widetilde{\alpha}_1$ and setting it to zero results in

$$\begin{bmatrix} \mathbf{M} & \mathbf{h}_2 \\ \mathbf{h}_2^* & 1 \end{bmatrix} \begin{bmatrix} \alpha_1^* \\ -\beta_2 \end{bmatrix} = \begin{bmatrix} \mathbf{h}_1 \\ 0 \end{bmatrix}, \qquad (25)$$

where $\mathbf{M} = \mathbf{H}\mathbf{H}^* + \dfrac{1}{\Gamma}\mathbf{I}$, Γ is the signal to noise ratio, $\mathbf{I}$ is the 4 by 4 identity matrix, $\mathbf{h}_1$ is the first column of $\mathbf{H}$, and $\mathbf{h}_2$ is the second column of $\mathbf{H}$. It follows that

$$\alpha_I = (\mathbf{M} - \mathbf{h}_2 \mathbf{h}_2^*)^{-1} \mathbf{h}_1 \quad \text{and} \quad \beta_2^* = \mathbf{h}_2^* (\mathbf{M} - \mathbf{h}_2 \mathbf{h}_2^*)^{-1} \mathbf{h}_1. \qquad (26)$$

It can be shown that

$$(\mathbf{M} - \mathbf{h}_2 \mathbf{h}_2^*)^{-1} = \mathbf{M}^{-1} + \frac{\mathbf{M}^{-1} \mathbf{h}_2 \mathbf{h}_2^* \mathbf{M}^{-1}}{1 - \mathbf{h}_2^* \mathbf{M}^{-1} \mathbf{h}_2}, \qquad (27)$$

which yields

$$\beta_2^* = \frac{\mathbf{h}_2^* \mathbf{M}^{-1} \mathbf{h}_1}{1 - \mathbf{h}_2^* \mathbf{M}^{-1} \mathbf{h}_2}. \qquad (28)$$

From the structure of the matrix $\mathbf{H}$ we can easily verify that $\mathbf{h}_1$ and $\mathbf{h}_2$ are orthogonal. Using this fact and the structure of the matrix $\mathbf{M}$, it can be shown that

$$\beta_2 = 0 \tag{29}$$

$$\alpha_1 = \mathbf{M}^{-1}\mathbf{h}_1. \tag{30}$$

Hence, the MMSE IC solution given by equations (29) and (30) will minimize the mean-squared error in c_1 without any regard to c_2. Considering the alternative cost function when β_2 is set to 1, a similar analysis leads to the conclusion that

$$\beta_1 = 0 \tag{31}$$

$$\alpha_2 = \mathbf{M}^{-1}\mathbf{h}_2 \tag{32}$$

In this case, the MMSE IC solution given by equations (31) and (32) will minimize the mean-squared error in c_2 without any regard to c_1. Therefore, from equations (29)-(32), we can easily see that the MMSE interference canceller for signals from terminal unit 10 will consist of two different sets of weights α_1 and α_2 for c_1 and c_2, respectively. The weights for decoding signals from terminal 30 can be obtained in a similar fashion, as expected. Thus, the decoding of signals from terminal units 10 and 30 can be performed with a single subroutine MMSE.DECODE in decoder 25 as follows:

$$(\mathbf{c}, \Delta_c) = \text{MMSE.DECODE}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{H}_1, \mathbf{H}_2, \mathbf{G}_1, \mathbf{G}_2, \Gamma)$$

{

$$\tilde{\mathbf{r}} = \begin{bmatrix} \mathbf{r}_1^T & \mathbf{r}_1^T \end{bmatrix}^T$$

$$\tilde{\mathbf{H}} = \begin{bmatrix} \mathbf{H}_1 & \mathbf{G}_1 \\ \mathbf{H}_2 & \mathbf{G}_2 \end{bmatrix}$$

$$\mathbf{M} = \mathbf{HH}^* + \frac{1}{\Gamma}\mathbf{I}$$

$$\mathbf{h}_1 = \begin{bmatrix} \mathbf{h}_{11} & \mathbf{h}_{21}^* & \mathbf{h}_{12} & \mathbf{h}_{22}^* \end{bmatrix}^T = \text{first column of } \mathbf{H}$$

$$\mathbf{h}_2 = \begin{bmatrix} \mathbf{h}_{21} & -\mathbf{h}_{11}^* & \mathbf{h}_{22} & -\mathbf{h}_{21}^* \end{bmatrix}^T = \text{second column of } \mathbf{H}$$

$$\alpha_1^* = \mathbf{M}^{-1}\mathbf{h}_1 \,, \alpha_2^* = \mathbf{M}^{-1}\mathbf{h}_2$$

$$\hat{\mathbf{c}} = \arg\min_{\hat{c}_1, \hat{c}_2 \in C} \{\|\alpha_1^*\tilde{\mathbf{r}} - \hat{c}_1\|^2 + \|\alpha_1^*\tilde{\mathbf{r}} - \hat{c}_2\|\}$$

$$\Delta_c = \|\alpha_1^*\tilde{\mathbf{r}} - \hat{c}_1\|^2 + \|\alpha_1^*\tilde{\mathbf{r}} - \hat{c}_2\|^2$$

}

With such a subroutine, both $\hat{\mathbf{s}}$ and $\hat{\mathbf{c}}$ can be estimated, as follows:

$$(\hat{\mathbf{c}}, \Delta) = \text{MMSE.DECODE}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{H}_1, \mathbf{H}_2, \mathbf{G}_1, \mathbf{G}_2, \Gamma) \tag{33}$$
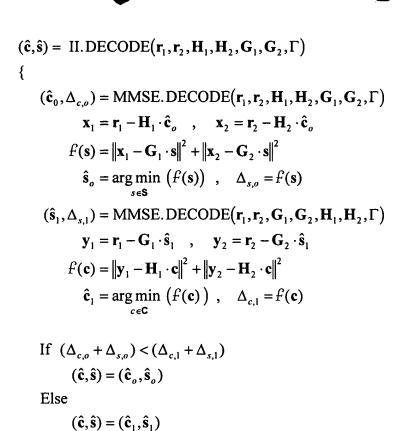
11

$$(\hat{\mathbf{s}},\Delta)=\text{MMSE.DECODE}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{G}_1,\mathbf{G}_2,\mathbf{H}_1,\mathbf{H}_2,\Gamma) \qquad (34)$$

Additional improvement can be realized by employing a two-step interference cancellation approach. In this two-step approach, the receiver decodes signals from both terminals using the subroutine MMSE.DECODE disclosed above. Assuming that symbols

5 from the terminal unit 10 $\hat{\mathbf{c}}_o$ have been decoded correctly, the receiver can, then, *perfectly* cancel the contribution of the terminal unit 10 in the received signal vectors $\mathbf{r}_1$ and $\mathbf{r}_2$. The receiver then uses $\mathbf{x}_1$ and $\mathbf{x}_2$, the received signal vectors after canceling signals from terminal unit 10, to *re-decode* symbols $\hat{\mathbf{s}}_o$ from terminal unit 30 using the optimum ML decoding rule in equation (10). Assuming that the symbols from terminal unit 10 have

10 been decoded correctly, the performance for terminal unit 30 will be equivalent to that with 2 transmit and 2 receive antennas (which is equivalent to 4 branch MRC diversity). The receiver then repeats the above step assuming that symbols from terminal unit 30 $\hat{\mathbf{s}}_1$ have been decoded correctly using the MMSE.DECODE subroutine. As before, the receiver cancels the contribution of terminal unit 30 in the received signal vectors $\mathbf{r}_1$ and uses $\mathbf{y}_1$

15 and $\mathbf{y}_2$, the received signal vectors after cancelling signals from terminal unit 30, *to re-decode* symbols from terminal unit 10 $\hat{\mathbf{c}}_1$ using the optimum ML decoding rule in equation (10). Also as before, assuming that symbols from terminal unit 30 have been decoded correctly, the performance for terminal unit 10 will be equivalent to that with 2 transmit and 2 receive antennas. Letting $\Delta_0 = \Delta_{c_0} + \Delta_{s_0}$ and $\Delta_1 = \Delta_{c_1} + \Delta_{s_1}$ denote the overall

20 uncertainty for $\hat{\mathbf{c}}_0$ and $\hat{\mathbf{s}}_0$, and for $\hat{\mathbf{c}}_1$ and $\hat{\mathbf{s}}_1$, respectively, the receiver compares the overall uncertainty and chooses the pair $(\hat{\mathbf{c}}_o,\hat{\mathbf{s}}_o)$ if $\Delta_o < \Delta_1$, and $(\hat{\mathbf{c}}_1,\hat{\mathbf{s}}_1)$ otherwise. The two-step interference cancellation and ML decoding algorithm is presented below in pseudo-code subroutine II.MMSE.DECODE.

$$(\hat{c}, \hat{s}) = \text{II.DECODE}(r_1, r_2, H_1, H_2, G_1, G_2, \Gamma)$$

$$\{$$

$$(\hat{c}_0, \Delta_{c,o}) = \text{MMSE.DECODE}(r_1, r_2, H_1, H_2, G_1, G_2, \Gamma)$$

$$x_1 = r_1 - H_1 \cdot \hat{c}_o \quad , \quad x_2 = r_2 - H_2 \cdot \hat{c}_o$$

$$f(s) = \|x_1 - G_1 \cdot s\|^2 + \|x_2 - G_2 \cdot s\|^2$$

$$\hat{s}_o = \arg\min_{s \in S} \left(f(s) \right) \quad , \quad \Delta_{s,o} = f(s)$$

$$(\hat{s}_1, \Delta_{s,1}) = \text{MMSE.DECODE}(r_1, r_2, G_1, G_2, H_1, H_2, \Gamma)$$

$$y_1 = r_1 - G_1 \cdot \hat{s}_1 \quad , \quad y_2 = r_2 - G_2 \cdot \hat{s}_1$$

$$f(c) = \|y_1 - H_1 \cdot c\|^2 + \|y_2 - H_2 \cdot c\|^2$$

$$\hat{c}_1 = \arg\min_{c \in C} \left(f(c) \right) \quad , \quad \Delta_{c,1} = f(c)$$

$$\text{If } (\Delta_{c,o} + \Delta_{s,o}) < (\Delta_{c,1} + \Delta_{s,1})$$

$$(\hat{c}, \hat{s}) = (\hat{c}_o, \hat{s}_o)$$

$$\text{Else}$$

$$(\hat{c}, \hat{s}) = (\hat{c}_1, \hat{s}_1)$$

$$\}$$

With appreciation of the above-disclosed theoretical background, we realized that enhanced performance is attained by devoting the space-time block coding to interference cancellation and ML decoding, while another coding schema may be used to overcome channel-caused degradations, such as fading. Accordingly, each of the transmitters in FIG. 1 includes a channel coder (14 and 34, respectively) that is interposed between the input signal and the transmitter's space-time block coder. Channel coders 14 and 24 can employ any conventional channel error correcting code (for example a trellis code, or a convolutional code).

At receiver 20, the inner space-time block code is decoded in element 26 and is used to suppress interference from the various co-channel terminals, using the MMSE approach disclosed above. Element 26 forms the two interference canceling vectors α_{i1} and α_{i1} corresponding to some terminal, i, and element 27 forms the two decision variables

$$\xi_1 = \alpha_1^* r \quad and \quad \xi_2 = \alpha_2^* r. \tag{35}$$

Those decisions, however, are used as soft decisions of the transmitted information

symbols, and are fed to channel decoder 28, which is a conventional decoder that corresponds to the type of encoding performed in channel encoders 14 and 34. Thus, in the arrangement depicted in FIG. 1, the structure of an inner coder is used for interference suppression, such that many co-channel terminals can operate simultaneously while

5   providing diversity. The output of the inner code space-time decoder forms the input to the outer coder decoder, which decides on the transmitted information while providing protection against channel errors.

FIG. 2 presents an arrangement for increasing the data rate or throughput in wireless systems. In FIG. 2, the information to be transmitted is demultiplexed in element

10   40 into two streams. One stream is applied to channel encoder 41, and the other stream is applied to channel encoder 51. The output of channel encoder 41 is applied to space-time block encoder 42, and then to mapper and pulse shaper 43 and antennas 44 and 45. Similarly, the output of channel encoder 51 is applied to space-time block encoder 52, and then to mapper and pulse shaper 53 and antennas 54 and 45. Generalizing, information

15   symbols from a transmitting terminal are split into L parallel streams. Stream l is then encoded using a channel code with rate R_l and then coded with a space-time block encoder with N transmitting antennas. The coding rates can be the same, but an advantage accrues when the coding rates are chosen such that $R_1 > R_2 > \cdots > R_L$. In such a case, symbols transmitted in stream l will have better immunity against channel errors as compared to

20   symbols transmitted in stream u where $u>l$. The base station receiver is assumed to be equipped with at least L receive antennas. The base station receiver treats each stream as a different user and uses the iterative interference cancellation techniques disclosed above, or the ones disclosed in the aforementioned '163 application. Since the first stream has the smallest coding rate R_l, it has the best immunity against the channel errors and most likely

25   it will be error free. The receiver then uses the decoded symbol of stream l to subtract the contributions of the first stream in the total received signals, while decoding the remaining L-l streams. In decoding the remaining L-l streams, the decoder decodes signals from the second stream first, since it has the best immunity against channel errors among the remaining L-l streams (due to its lowest rate, R_2 from among the remaining streams).

30   Then the receiver uses the decoded symbols for the second stream to cancel out its contribution in the received signal. This process is repeated until all streams are decoded.

14

It can be shown that, in this case, the system throughput is given by

$$\rho = \frac{1}{L} \sum_{l=1}^{L} R_l (1 - FER_l),$$ (36)

where FER_l is the frame error rate of stream l.